Comparison of photonic reservoir computing systems for fiber transmission equalization

A. Argyris, J. Cantero, M. Galletero, E. Pereda, C. R. Mirasso, I. Fischer, and M. C. Soriano, *Senior Member, IEEE*

***Abstract*—In recent years, various methods, architectures, and implementations have been proposed to realize hardware-based reservoir computing (RC) for a range of classification and prediction tasks. Here we compare two photonic platforms that owe their computational nonlinearity to an optically injected semiconductor laser and to the optical transmission function of a Mach-Zehnder modulator, respectively. We numerically compare these platforms in a delay-based reservoir computing framework, in particular exploring their ability to perform equalization tasks on nonlinearly distorted signals at the output of a fiber-optic transmission line. Although the non-linear processing provided by the two systems is different, both produce a substantial reduction of the bit-error-rate (BER) for such signals of up to several orders of magnitude. We show that the obtained equalization performance depends significantly on the operating conditions of the physical systems, the size of the reservoir and the output layer training method.**

I. Introduction

Reservoir computing (RC) is a machine learning method with a single hidden layer that is capable of processing sequential data streams [1]. RC is based on networks with recurrent connections and relies on mapping an input signal onto a high dimensional state space to facilitate tasks such as temporal pattern classification or nonlinear prediction. The training in RC is solely performed on the weights of the readout layer, while those of the recurrent network are kept fixed [2], [3]. It has been shown that certain RC configurations are suitable for hardware-implementations on physical substrates, see for a review [4], [5]. In particular, RC based on photonic components have been successfully used to solve various classification tasks with high speed and predicting complex timeseries with an extended prediction horizon. Moreover, photonic RC implementations have been recently demonstrated as alternative to digital signal processing solutions for signal recovery in optical transmission systems [6]–[8]. Photonic RC implementations that have been introduced so far are based on delay-coupled semiconductor lasers and semiconductor optical amplifiers [6], [9]–[12], optoelectronic delay systems [13]–[17], passive photonic elements [18], photonic crystal cavities [19] and silicon photonic chips [20], [21], or fiber nonlinearities [7], [22] as the photonic nonlinear elements for computation, see for a review [23].

Various implementations with differing nonlinearities exhibit good performance, and it has not been clear so far if there is any nonlinearity that outperforms the others when used for a certain RC task. Some desired reservoir attributes may be determined by the nature of the tasks that need to be solved. For example, one would rationally exploit reservoir schemes with enhanced memory properties to process data sequences with a significant inherent dependence on their previous states. In addition, one could also exploit nonlinear properties of the components for pattern-discrimination tasks.

Here we compare two delay-based photonic RC systems of different nature; one based on a semiconductor laser nonlinearity and the other based on an Ikeda nonlinearity realized optoelectronically. We deploy these two systems in a different way regarding the input information rate. For the first system, we consider a fast modulation input rate that induces nonlinear transients, while for the second system we consider a slow modulation input rate that excludes any transient operation of the nonlinearity. Both systems are evaluated on their ability to process data sequences that undergo a nonlinear transformation with memory effects when sent through a nonlinear fiber transmission channel.

II. Delay-based Reservoir Computing

RC follows a typical architecture of input layer, hidden layer (reservoir), and output layer. The input layer contains the input S of dimension D. In the traditional RC configuration, the reservoir, or hidden layer, contains N nonlinear nodes, often connected in a random fashion described by the reservoir connectivity matrix W. This reservoir connectivity matrix, of dimension NxN, allows for recurrent connections within the reservoir. The information to be processed S is injected into the reservoir via an input connectivity matrix W_{in} of dimension DxN which is also often chosen randomly. The reservoir nodes are connected to the output layer with connectivity matrix W_{out}, which has a dimension NxO (when the output layer is of dimension O). In RC, only the output weights W_{out} are optimized using a supervised learning procedure [24]. In a

This work was supported by MINECO (Spain), through project TEC2016-80063-C3 (AEI/FEDER, UE). We acknowledge the Spanish State Research Agency, through the Severo Ochoa and María de Maeztu Program for Centers and Units of Excellence in R&D (MDM-2017-0711). The work of MCS was supported by the Spanish Ministerio de Economía, Industria y Competitividad through a Ramon y Cajal Fellowship (RYC-2015-18140). AA was supported by the Conselleria d'Innovació, Recerca i Turisme del Govern de les Illes Balears and the European Social Fund.

AA, JC, MG, CM, IF and MS are with the Instituto de Física Interdisciplinar y Sistemas Complejos, IFISC (CSIC-UIB), Campus Universitat Illes Balears, E-07122 Palma de Mallorca, Spain e-mail: (see http://ifisc.uib-csic.es/people).

EP is with the Department of Industrial Engineering & Institute of Biomedical Technology, University of La Laguna, Tenerife, Spain and the Laboratory of Cognitive and Computational Neuroscience, Center for Biomedical Technology, University Polytechnic of Madrid, Madrid, Spain.

Manuscript received April 1, 2019; revised xxxx.

general form, the dynamical response of the reservoir X can be written as:

$$\dot{X}[t_k] = f(WX[t_k] + W_{in}S[t_k]), \tag{1}$$

where f describes the details of the dynamical system that can contain nonlinearities, delay, etc. After training the output weights, the readout nodes of the reservoir computer are at a given time t_k defined as:

$$Y[t_k] = W_{out}X[t_k]. \tag{2}$$

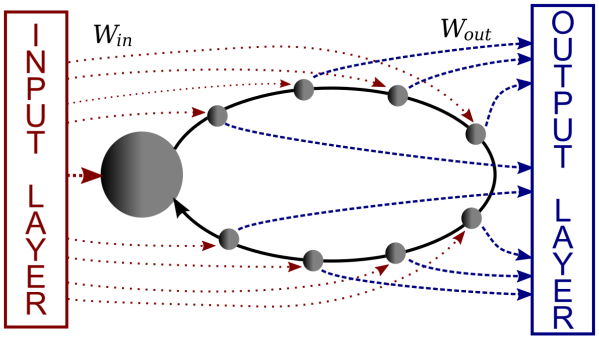

Fig. 1. (color online). RC based on a single nonlinear node with delay.

A significant simplification of the RC concept is to consider a ring connectivity in the reservoir [25]. The ring connectivity can in turn be implemented with a single nonlinear node and a delayed feedback loop allowing to generate multiple virtual copies of the nonlinear node via time-multiplexing [26], [27], simplifying even further the RC concept for hardware implementations. This delay-based RC structure is illustrated in Fig. 1. The virtual nodes are usually taken at equidistant temporal positions within the response of the nonlinear node.

In our work, we consider an input connectivity matrix W_{in} with values drawn from a Gaussian distribution of zero mean and unit standard deviation.

III. Photonic Systems

A. Semiconductor laser with time-delayed optical feedback

Semiconductor lasers subject to optical feedback are systems that have been investigated extensively during the last decades (see, e.g., [28] and references therein). The variety of dynamical emission properties that they exhibit established them as unique building blocks for modern technological applications, such as communication encryption [29], rainbow refractometry [30], random number generation [31], fast signal processing [10] and more. The Lang-Kobayashi rate equations model effectively describes the behavior of this system for various conditions, even when considering additional external perturbations [32]. Here, we adapt this model to describe the operation of a photonic reservoir, based on a semiconductor laser subject to optical feedback with time delay τ_{SL}, that processes the external time-dependent information stream in Eqs. 3-6:

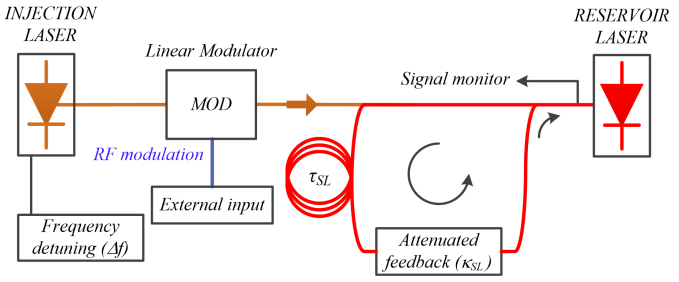

Fig. 2. (color online). Simplified scheme of a photonic reservoir with a SL with optical feedback.

$$\begin{aligned}
\dot{E}(t) &= \frac{1}{2}(1+i\alpha)\left(G-\gamma_{ph}\right)E(t) + \kappa_{SL}\gamma_{in}E(t-\tau) \\
&+ \kappa_{inj}\gamma_{in}E_{inj}(t)e^{i2\pi\Delta f t}, && (3)\\
\dot{n}(t) &= \frac{J_{SL}(t)}{e} - \gamma_e n(t) - G\left|E(t)\right|^2, && (4)\\
G &= g_n(n(t)-n_0)(1+\epsilon\left|E(t)\right|^2)^{-1}, && (5)\\
E_{inj}(t) &= E_0(1+\gamma'_{SL}W_{in}S(t)), && (6)
\end{aligned}$$

The conceptual scheme for this photonic RC system is depicted in Fig. 2. The dynamical behavior of the reservoir's laser is defined by the feedback rate κ_{SL}. The external input is injected into the reservoir via an independent optical carrier with frequency detuning Δf with respect to the reservoir's laser emission. The average injected optical field is determined by the values of κ_{inj} and E_0, which are selected to avoid complete injection locking of the reservoir laser to the injected signal [11]. The external input to be processed by the reservoir ($W_{in}S(t)$) is then applied as a linear modulation signal onto the injected optical field, with a modulation index $\gamma'_{SL} = \gamma_{SL}(\sqrt{D})^{-1}$ that incorporates also a normalization on the input dimension D. A summary of the parameter values used in the numerical simulations is listed in Tab. I. The chosen number of virtual nodes defined in the reservoir's time delay loop determines the feedback delay τ_{SL}, while the temporal spacing between the virtual nodes is set to 25 ps. This small spacing is still adequate to establish a coupling due to inertia between neighboring virtual nodes, since the characteristic frequency of the reservoir system (with optical injection and feedback) is calculated to be $\sim$ 8.6 GHz. Thus, the 25 ps time separation between the virtual nodes of the reservoir is lower than the characteristic time by a factor of 4.6, close to the factor 5 that is commonly used in such delay-based RC systems [26]. Unless otherwise stated in some parts of this work, the number of virtual nodes defined in the examined reservoirs is set to $N = 400$, resulting in a feedback-delay time of 10 ns.

Note that we have neglected any noise terms for the laser operation and signal detection at the reservoir output. This is because our aim here is to evaluate the reservoir computational efficiency rather than its performance in a noisy environment.

TABLE I
PARAMETER VALUES USED IN THE NUMERICAL SIMULATIONS OF A SEMICONDUCTOR LASER SUBJECT TO OPTICAL FEEDBACK.

Parameter	Value
α	3
γ_{ph}	500 ns^{-1}
κ_{SL}	*variable*
κ_{inj}	0.4
τ_{SL}	*variable*
Δf	*variable*
γ_e	0.5 ns^{-1}
γ_{in}	100 ns^{-1}
g_n	1.2·10^{-5} ns^{-1}
ϵ	5·10^{-7}
λ	1.55 μm
n_0	1.5·10^{8}
e	1.602·10^{-19} Cb
J_{SL}	15.3 mA
$J_{SL,thr}$	15.37 mA
γ_{SL}	*variable*
E_0	100

B. Optoelectronic nonlinearity with time-delayed optical feedback

K. Ikeda originally proposed a nonlinear optical system with delayed feedback to demonstrate the existence of optical chaos [33]. Since then, the Ikeda system has become a paradigmatic model to investigate the dynamical complexity of delay systems in optics [34], [35]. The model is based on a delay differential equation with a sine nonlinearity and a delayed feedback term [36]. Here, we use an optoelectronic implementation of the Ikeda system in the map limit as a second scheme to implement a delay-based photonic reservoir computer. Figure 3 shows the conceptual scheme of the system, which includes a laser source, a Mach-Zehnder modulator providing a nonlinear transmission function and an optical fiber for the delay loop. When driven by an external time-dependent information stream, the temporal evolution in the Ikeda map description reads [15]:

$$z[k] = \sin^2\left(\kappa_{oe} z[k-1] + \gamma'_{oe} W_{in} S[k] + \phi\right), \qquad (7)$$

where κ_{oe} controls the feedback strength, $\gamma'_{oe} = \gamma_{oe}(\sqrt{D})^{-1}$ is the input scaling with a normalization on the input dimension D, and ϕ is the phase of the system. The delay time τ_{oe} has been normalized to 1. Equation 7 can be used to implement a reservoir computer with a unidirectional ring connectivity when the value of the delay time is mismatched with the number of virtual nodes, which can be easily done by adapting the time-multiplexed input [14]. Here, the delay loop is defined as being one virtual node longer $(N+1)$ than the number of virtual nodes defined for the reservoir (N). In this manner, a virtual node with index i (i in the range from 1 to N) is connected to virtual node with index $i+1$ (modulo N) after the delay loop.

The map model in Eq. 7 assumes that the response of the nonlinearity is instantaneous, i.e., the response of the Mach-Zehnder modulator and all other components is much faster than the input information rate. This is in stark contrast to the semiconductor laser model in Eqs. 3-6, in which we consider that the input information rate is faster than the relaxation oscillation frequency of the reservoir's laser. From the point of view of an experimental realization, this means that the semiconductor laser reservoir operates in a transient regime while the optoelectronic system does not. We have chosen these different modelling approaches for the two photonic implementations, as these are the most common forms in which each of these systems have been operated, see for the semiconductor laser system in experiment [10], [11], [37] and modelling [38]–[40], and [14]–[16] for the optoelectronic system. Consequently, the virtual nodes in the SL case are coupled due to the system's inertia while they are coupled via a mismatch between masking period and feedback delay in the optoelectronic RC system. This results in having virtual reservoirs with different network topologies.

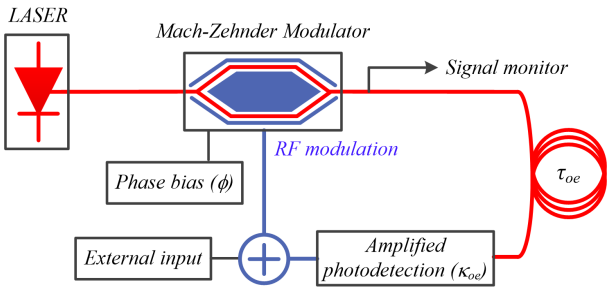

Fig. 3. (color online). Simplified scheme of an optoelectronic RC system with a sin^2 nonlinearity, given by a Mach-Zehnder modulator, and time-delayed feedback, provided by an optical fiber loop.

IV. RESULTS AND DISCUSSION

In the following, we compare the performance of the two optical systems for the same nonlinear channel equalization task. We focus on processing signals that emerge from an optical fiber communication channel. Specifically, the signals to be processed are distorted considerably, after 50 km of unamplified, uncompensated transmission at 1550nm and at 25 Gb/s bit rate. The numerical model of the transmission system that is used to generate these signals is presented in the Appendix, along with the definition of the relevant parameters. Each bit is represented by a pattern of 8 samples. The original and the degraded signal after 50 km of fiber transmission are shown in Fig. 4. The latter suffers from both chromatic dispersion and self-phase modulation effects. These effects result in mixing of neighboring bit patterns which eventually might destroy the possibility to recover the initial information. However, the attributes of the neighboring bit patterns can give useful information to recover the initial bit stream. Thus, to retrieve the present bit value we may consider the patterns of one (the present), three (the previous, the present and the next) or more bits as input to our processing task. The training at the output layer of the reservoir is then performed by considering the reservoir response in the timeframe corresponding to each bit pattern. By considering the difference between one bit time duration (40 ps) and the processing time of one bit in the case

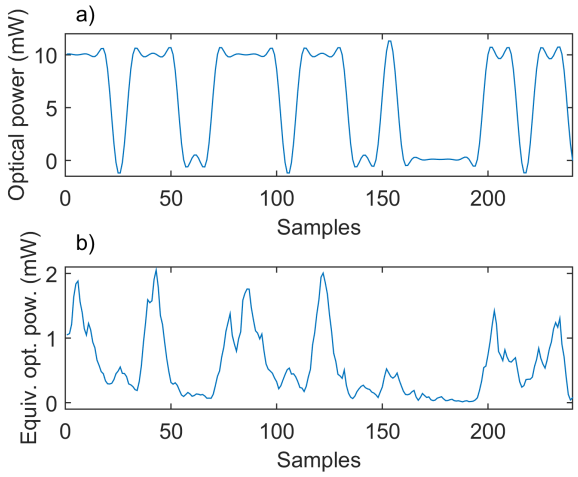

Fig. 4. (color online). Illustration of the optical input signal (a) before and (b) after 50km fiber transmission with a 25dB signal-to-noise ratio detected signal. Each bit is represented by 8 samples. Equiv. opt. pow. stands for equivalent optical power after detection.

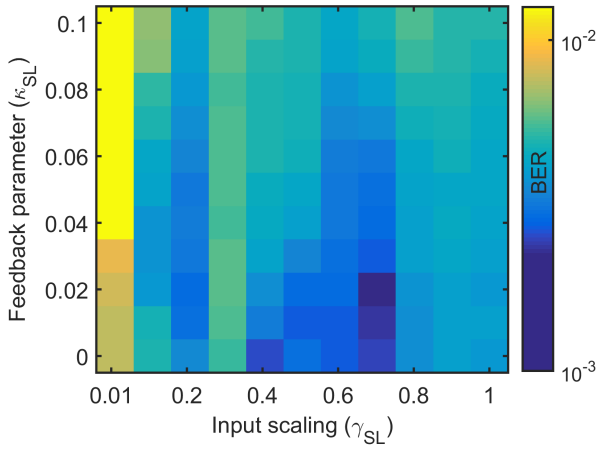

Fig. 5. (color online). BER as a function of the parameters κ_{SL} and γ_{SL}. The frequency detuning has been set to Δf=0GHz. Other parameter values as in Tab. I.

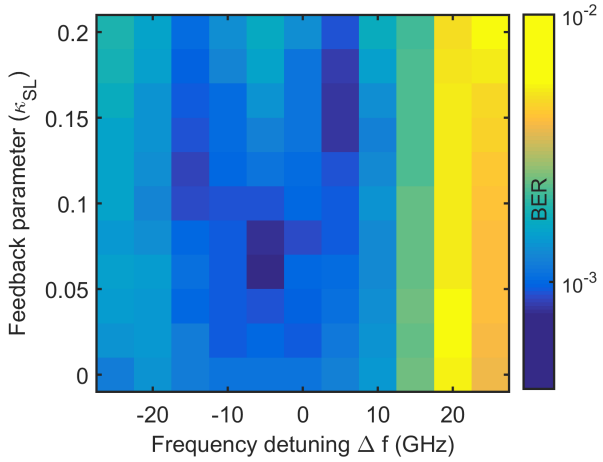

Fig. 6. (color online). BER as a function of the parameters κ_{SL} and Δf. The input scaling has been set to the optimum value γ_{SL}=0.7. Other parameters as in Fig. 5 and Tab. I.

of the semiconductor laser RC (one round-trip of the reservoir loop, so 10 ns time duration), we can determine the speed penalty of this processing method. In this case, the obtained penalty amounts to a factor of 250. Nevertheless, this value can be further reduced when considering shorter reservoirs with smaller number of nodes and alternative processing schemes [6]. The figure of merit for the overall RC performance of the equalization task is the bit error rate (BER) of the tested sequences, defined as the ratio of the erroneously recovered bits over the total number of bits.

When applying linear classification methods directly on the input signal and considering different extent of neighboring bit patterns as input, the BER of the recovered signal always remained above 0.04. We start by considering an input bit pattern of three bits length for the RC systems. Subsequently, the influence of the length of the neighboring bit patterns will also be explored, evaluating to what extent it can improve the computational performance.

The first step to determine efficient regimes for the RC performance is to optimize the parameters of Eqs. 3-6 and Eq. 7 for the given processing task. For this parameter optimization procedure, we consider a reservoir of size $N = 400$, an input block of 3 bits described by 8 samples each ($D = 3\text{x}8 = 24$), and two independent sequences of 16384 bits; one is used to train the system and the other one is used to evaluate the training efficiency (test sequence). In this case, the connectivity matrix W_{in} has a dimension of $D\text{x}N$ and is selected in all our investigations to be a random matrix, as previously discussed. As a result, the masked input $W_{in}S$ has a dimension N that can be sequentially fed to the N virtual nodes of the reservoir. The size of the train and test sequences is sufficient to obtain a well-trained reservoir system when using linear regression to compute the output weights W_{out}.

For cases in which low BER values were obtained (e.g. $< 10^{-3}$), extended sequences of 131072 bits are used to ensure an adequate resolution of the BER. For the semiconductor laser configuration, the required reservoir size results in a time-delayed feedback loop of τ_{SL} = 10 ns. We choose the parameter values for the model as provided in Tab. I and explore the dependence of the BER on the feedback parameter (κ_{SL}), the input scaling (γ_{SL}) and the frequency detuning (Δf). For the optoelectronic system, we evaluate the dependence of the BER on the feedback parameter (κ_{oe}), the input scaling (γ_{oe}) and the offset phase (ϕ). The parameters that have been varied for the two systems are those that had been identified to be the most critical accessible operational parameters of the reservoir that can be tuned.

Figures 5 and 6 show the BER dependency on the above parameters for the semiconductor laser based RC. The obtained results emerge after considering three different random input connectivity matrices W_{in} and evaluating the average RC performance on independent test data sets. To start, we make an attempt to identify an optimal input scaling value γ_{SL} for the processed information, by choosing zero frequency detuning between the injected optical field and the reservoir's

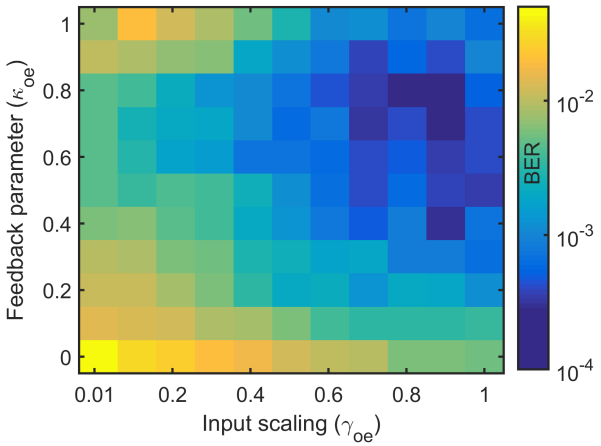

Fig. 7. (color online). BER as a function of the parameters κ_{oe} and γ_{oe}. The offset phase has been set to $\phi = 0.1\pi$.

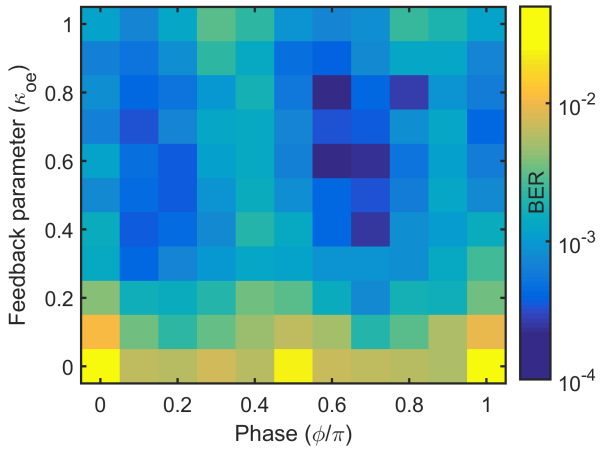

Fig. 8. (color online). BER as a function of the parameters κ_{oe} and ϕ. The input scaling has been set to the optimum value γ_{oe}=0.9.

laser optical field (Δf=0GHz). From the mapping in Fig. 5, we obtain a region of low BER values at intermediate values of the input scaling and at low values of the feedback parameter. Based on these results, we fix the value of the input scaling equal to γ_{SL}=0.7 and we follow an optimization process versus the feedback and the frequency detuning parameters. From the BER performance obtained in Fig. 6, we identify a region of low BER values at small frequency negative frequency detuning and moderate optical feedback values. In agreement with [11], this region corresponds to the boundaries of optical injection locking for the chosen laser parameters. From these results, we further select the optimum parameter values for κ_{SL}=0.06 and Δf=-5GHz that yield an average BER $\sim 7 \cdot 10^{-4}$. Nevertheless, since the semiconductor laser system is described by an extended parameter space, an even more efficient performance should not be excluded when considering a more detailed parameter optimization process.

Figures 7 and 8 illustrate the BER for the optoelectronic RC system as a function of the parameters in Eq. 7. These results are the average over three training and testing procedures with three different random input connectivity matrices W_{in}. In Fig. 7, we find a region of low BER (blue color) for parameter values in the range κ_{oe}=0.4-0.9 and γ_{oe}=0.7-0.9. Based on these results, we fixed the value of the input scaling to γ_{oe}=0.9. Subsequently, we proceeded to evaluate the parameter dependencies for the feedback parameter and the offset phase. Fig. 8 illustrates that several combinations of parameter values yield a performance close to the minimum obtained BER. Specifically, this minimum of the BER is $\sim 2{\cdot}10^{-4}$ and it is obtained for the parameters γ_{oe}=0.9, κ_{oe}=0.6, and $\phi = 0.6\pi$.

The selected set of the parameter values for the two photonic systems are summarized in Tab. II. Interestingly, we find a similarly low BER for the two photonic systems.

Beyond the performance results, it is possible to compare the two systems in more detail by looking at the nonlinear properties of the corresponding operating points. In both delay-based reservoir computing concepts illustrated in Fig. 1, we considered a single physical nonlinear node, namely a semiconductor laser or a Mach-Zehnder modulator. In addition

TABLE II
SELECTED PARAMETER VALUES USED IN THE NUMERICAL EVALUATION OF THE TWO PHOTONIC RESERVOIR SYSTEMS.

Parameter	Value
κ_{SL}	0.06
Δf	-5GHz
γ_{SL}	0.7
κ_{oe}	0.6
ϕ	0.6π
γ_{oe}	0.9

to the physical nonlinear node, we defined the virtual nodes as equidistant positions within the delay loop. Each of these virtual nodes is described by a different input transformation, due to a different input connectivity and the differing context of the neighbouring virtual nodes. The optimum shape of the nonlinear response for each virtual node may depend on the chosen benchmark task. Consequently, we focus on the diversity in the responses for each virtual node and compare these properties for the two photonic systems.

Figure 9 shows the responses for four arbitrarily selected virtual nodes in each photonic reservoir system. Specifically, in Fig. 9(a)-(d), we show the nonlinear responses for four virtual nodes for optimum operating conditions for the semiconductor-laser RC system. The nonlinear dynamics of semiconductor lasers subject to optical injection and delayed optical feedback is particularly rich. We find that different virtual nodes generate a qualitatively different nonlinear transformations of the input. Figure 9 (e)-(h) illustrates the nonlinear response of four different virtual nodes for the optoelectronic RC system for the parameter values in Tab. II. In this case, for all virtual nodes depicted in Fig. 9 (e)-(h), the characteristic scalar sin^2-shaped nonlinearity can clearly be recognized. Nevertheless, the responses of the virtual nodes exhibit sufficient differences in the covered range of the nonlinearity. The apparent similarity between the nonlinear transformations in the optoelectronic RC system does not hinder good computational performance. We find that, with

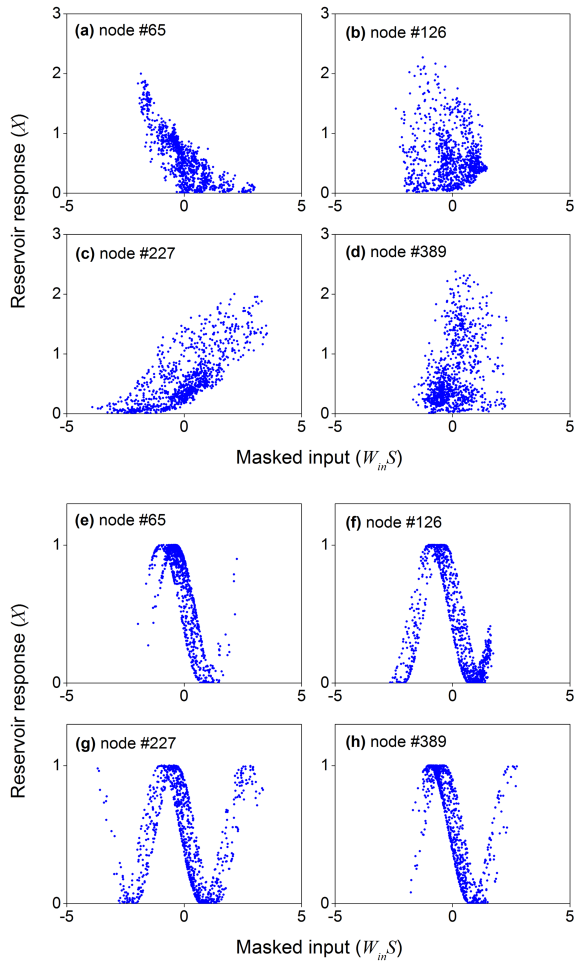

Fig. 9. (color online). Exemplary reservoir states for the two photonic systems. (a)-(d) Semiconductor laser. (e)-(h) Optoelectronic system with an Ikeda nonlinearity.

these reservoir node states, the output layer is still able to approximate the desired target bits.

As a further comparison of the performance of the two photonic systems, we evaluate the dependence of the BER on the length of the input pattern. This is done by varying the number of input bits considered at the RC input layer and adapting accordingly the dimensions of the input connectivity matrix. We always keep the central bit timeframe as the target of the RC training procedure and the reservoir response associated to the input pattern as the training output. Figure 10 depicts the BER depending on the number of input bits for both systems. We find that for both systems the minimum BER is obtained when the number of input bits is 3. The two photonic systems yield comparable results for 5 and 7 input bits, while the optoelectronic system yields lower errors for 1 and 3 input bits. Observing the distorted bit patterns in Fig. 4, it is relatively unexpected that we find the minimum BER for an input pattern of only 3 bits length. To understand this dependence of the BER on the length of the input pattern in more detail, we compare it to the BER obtained when applying linear regression directly to the input patterns (without going through the nonlinear reservoir). Figure 10 (blue line) shows

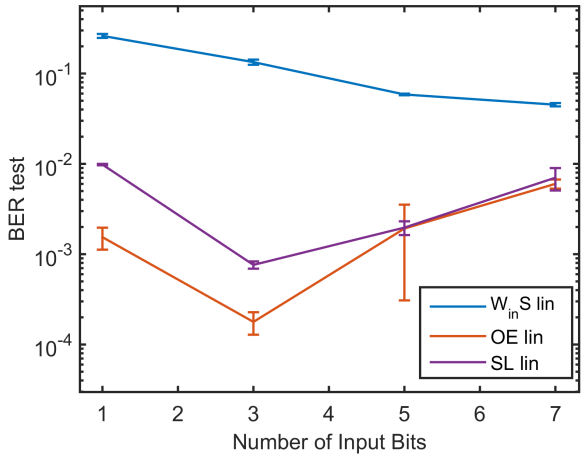

Fig. 10. (color online). BER as a function of the number of bits considered at the input layer. The output weights have been computed via linear regression (lin). The number of virtual nodes is $N = 400$ and other parameters are as in Tabs. I and II. OE stands for optoelectronic, SL for semiconductor laser and $W_{in}S$ for masked input.

that in the latter case the BER shows a decreasing trend for an increasing length of the input pattern, saturating for larger input patterns. Since one of the main advantages of reservoir computing systems is the ability to exploit its intrinsic memory of past inputs, we can attribute the minimum BER for 3 input bits to the optimum combination of the fading memory of the reservoir and the properties of the input stream. Using the BER computed directly over the input data as reference, we can assess the improvement yielded by the nonlinear photonic RC system. We find that this improvement is well beyond one order of magnitude for the chosen combination of parameters.

Finally, we explore the dependence of the BER on the size of the reservoir, i.e. the number of virtual nodes N, for the two photonic systems. We show in Fig. 11 that, for both systems, the BER decreases strongly when increasing the size of the reservoir. For the RC approach with output weights computed via linear regression, the BER reaches our statistical significance level ($7.6{\cdot}10^{-6}$, test sequence of 131072 bits) for a reservoir size of $N = 800$. We find that the BER of the two photonic RC systems exhibit a similar decreasing trend when the size of the reservoir increases.

Since the recovery of the transmitted bits can be seen as a dynamical pattern classification task, we tested whether an optimization procedure using a logistic regression [41], instead of a linear one, would yield even lower errors. As shown in Fig. 11 this is indeed the case. The BER is further reduced when the output weights are computed via logistic regression. For the particular cases of $N = 200$ and $N = 400$, we find a BER improvement of one to two orders of magnitude when the output layer is optimized with the logistic regression over the linear one. The technical details associated with the computation of the logistic regression can be found in Appendix B. For this particular equalization task, the optoelectronic RC system with Ikeda nonlinearity tends to yield a lower BER than the semiconductor laser nonlinearity, but both systems eventually reach the statistical significance level for a sufficiently large

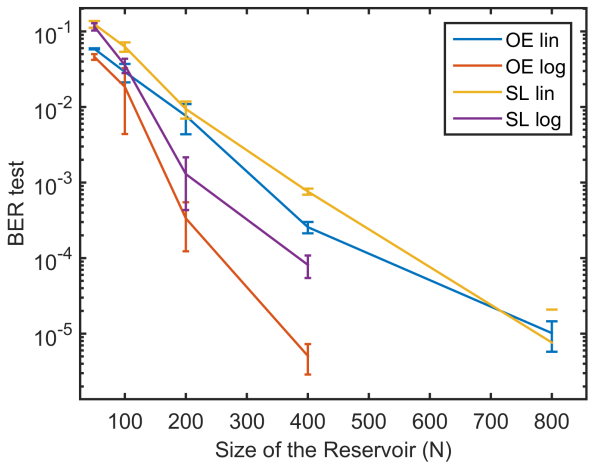

Fig. 11. (color online). BER as a function of the number of reservoir nodes. The output weights have been computed via a linear regression (lin) or a logistic optimization (log). Note that for the logistic regression and $N \geq 800$, we obtain error rates below the $7.6{\cdot}10^{-6}$ resolution, for both systems.

number of virtual nodes. The lower BER of the RC system based on the Ikeda map (Eq. 7) may be related to having a higher computational capacity than the RC system based on the SL model (3-6). Since the computational capacity is given by the number of independent virtual nodes and several neighboring virtual nodes are correlated due to inertia in the semiconductor laser based RC, there are indications that the capacity of the semiconductor laser based RC is smaller than that of the optoelectronic RC system for the same number of virtual nodes. Nevertheless, using inertia coupling allows to better exploit the bandwidth of the reservoir, therefore allowing for higher spectral efficiency.

V. Conclusion

We compared two delay-based photonic RC systems that are different in two important aspects. First, we considered RC systems based either on a semiconductor laser nonlinearity or on an optolectronic system with Ikeda nonlinearity, which are clearly of a different nature. Second, we drove the systems in two different ways regarding the input information rate. We considered a fast modulation input rate, that induces nonlinear transients, in the case of the semiconductor laser. In contrast, we considered a slow modulation input rate in the case of the optoelectronic system that can also be interpreted as an instantaneous response of the Mach-Zehnder sin^2 modulator. Despite these apparent differences, the performances of the two photonic RC systems go hand in hand. If one concentrates on the similarities between the two RC systems, we identify that the size of the virtual reservoir (N) is a crucial parameter that greatly affects the ultimate performance of the system.

Regarding the performance for the fiber transmission equalization, we find that in both models we need to accurately tune the parameters to find the respective optimum accuracy. By doing so, we reach the statistical significance level of the BER ($7.6{\cdot}10^{-6}$) for the two photonic RC systems. In particular, we find a clear improvement over the reference BER obtained from applying linear regression directly to the input data ($\sim 4{\cdot}10^{-2}$). To gain more insight into the properties of the optimum operating points, we have visualized the nonlinear responses of the virtual nodes in each of the systems. While we find that a diversity of the nonlinear responses is an attribute of the optimum operating points, we still lack a quantitative indicator explaining why some operating points are better than others. Finding such indicators represents a general challenge in machine learning methods, since the optimum operating point depends on the properties of the method as well as on the input and the desired task.

Based on our comparison of the two different photonic RC approaches, we can state that both of them are promising candidates for achieving fiber transmission equalization using hardware-implemented photonic systems. We note that for practical hardware implementations further aspects like signal detection and noisy environments will also play a role, albeit neglected here. The consideration of these additional aspects might reduce the final margin of BER improvement in hardware-implemented RC, but our approach still has been demonstrated to show comparable improvements in practice [6].

Appendix

A. Signal propagation

The input signal is obtained numerically using the nonlinear Schrödinger equation (NLSE) propagation model for single mode fiber transmission. We focus on single-channel, single-polarization transmission considering all relevant effects that affect signal propagation, such as optical attenuation, chromatic dispersion and Kerr nonlinearity. The optical peak power launched for transmission is equal to 10 mW and is emitted from a distributed feedback laser, with relative intensity noise of -153 dB/Hz. This power level is not considered high enough to trigger stimulated Brillouin scattering nonlinear effects, given the modulation bandwidth of the baseband 25 Gb/s two-level pulse amplitude modulation encoding. Thus, the nonlinear Schrödinger equation is written in a form that includes only relevant phenomena to our investigation. The slowly varying optical field $E_{tr}(t,z)$ that travels along the fiber is described by:

$$i\frac{\partial E_{tr}}{\partial t} + i\frac{\alpha_{loss}}{2}E_{tr} - \frac{\beta_2}{2}\frac{\partial^2 E_{tr}}{\partial t^2} + \gamma|E_{tr}|^2 E_{tr} = 0 \qquad (8)$$

where z = 50 km is the transmission distance, t is the relative time in the frame that moves with the envelope velocity, α_{loss} = 0.2 db/km is the fiber transmission loss coefficient, β_2 = 17 ps/nm/km is the chromatic dispersion coefficient and γ = 2.7$\cdot$ 10^{-20} m^2/W is the Kerr nonlinear coefficient, at 1550 nm. The optical signal after transmission is photodetected with a typical PIN photoreceiver, with responsivity of 0.9 A/W, in a direct-detection scheme. Thermal and dark current noise are included in the photodetection stage, while the bandwidth response of the detection has a low cutoff frequency of 0.8 the data bit-rate. Finally, the detected signal is obtained with a signal-to-noise ratio of 25 dB and is used as the external input that is fed into the reservoir (see Figs. 2 and 3).

B. Linear and logistic regression

Many optimization problems involve computing the minimization of a fitting function over a large number of data points. In the case of a linear regression, the least-squares method is the simplest and most common way to solve the problem of finding the best fitting straight line through a set of points. The function to be minimized in linear regression can be defined as:

$$\underset{w\in\mathbb{R}^p}{\text{minimize}}\frac{1}{n}\sum_{i=1}^{n}(x_i^T w - y_i)^2, \tag{9}$$

where $x_i \in \mathbb{R}^p$ and $y_i \in \mathbb{R}$ are the data sample pairs (i.e. inputs and targets) associated with a regression problem.

Logistic regression is a widely used method in machine learning for binary classification. Unlike linear regression which outputs continuous number values, logistic regression transforms its output using the logistic function to return a probability value which can then be mapped to two (or eventually more) discrete classes. The corresponding function to be minimized in the case of logistic regression can be defined as [41]:

$$\underset{w\in\mathbb{R}^p}{\text{minimize}}\frac{1}{n}\sum_{i=1}^{n}\log(1+\exp(-y_i x_i^T w)), \tag{10}$$

where $x_i \in \mathbb{R}^p$ and $y_i \in \{-1, 1\}$ are the data sample pairs associated with a binary classification problem.

In this work, we have computed the logistic regression using the stochastic average gradient (SAG) algorithm [41]. This method is efficient both in terms of iteration costs and convergence rates. In general, stochastic gradient methods benefit from the fact that there is often a large amount of redundancy between examples when the number of training examples n is very large.

REFERENCES

[1] D. Verstraeten, B. Schrauwen, M. d'Haene, and D. Stroobandt, "An experimental unification of reservoir computing methods," *Neural networks*, vol. 20, no. 3, pp. 391–403, 2007.

[2] W. Maass, T. Natschläger, and H. Markram, "Real-time computing without stable states: A new framework for neural computation based on perturbations," *Neural Computation*, vol. 14, no. 11, p. 2531, 2002.

[3] H. Jaeger and H. Haas, "Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication," *Science*, vol. 304, p. 78, 2004.

[4] M. Dale, J. F. Miller, and S. Stepney, *Reservoir Computing as a Model for In-Materio Computing*. Cham: Springer International Publishing, 2017, pp. 533–571.

[5] G. Tanaka, T. Yamane, J. B. Héroux, R. Nakane, N. Kanazawa, S. Takeda, H. Numata, D. Nakano, and A. Hirose, "Recent advances in physical reservoir computing: a review," *Neural Networks*, 2019.

[6] A. Argyris, J. Bueno, and I. Fischer, "Photonic machine learning implementation for signal recovery in optical communications," *Scientific Reports*, vol. 8, no. 1, p. 8487, 2018.

[7] M. Sorokina, S. Sergeyev, and S. Turitsyn, "Fiber echo state network analogue for high-bandwidth dual-quadrature signal processing," *Optics express*, vol. 27, no. 3, pp. 2387–2395, 2019.

[8] A. Argyris, J. Bueno, and I. Fischer, "Pam-4 transmission at 1550nm using photonic reservoir computing post-processing," *IEEE Access*, vol. 7, pp. 37 017–37 025, 2019.

[9] F. Duport, B. Schneider, A. Smerieri, M. Haelterman, and S. Massar, "All-optical reservoir computing," *Optics express*, vol. 20, no. 20, pp. 22 783–22 795, 2012.

[10] D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer, "Parallel photonic information processing at gigabyte per second data rates using transient states," *Nat. Commun.*, vol. 4, p. 1364, 2013.

[11] J. Bueno, D. Brunner, M. C. Soriano, and I. Fischer, "Conditions for reservoir computing performance using semiconductor lasers with delayed optical feedback," *Opt. Express*, vol. 25, no. 3, pp. 2401–2412, Feb 2017.

[12] R. M. Nguimdo, E. Lacot, O. Jacquin, O. Hugon, G. Van der Sande, and H. G. de Chatellus, "Prediction performance of reservoir computing systems based on a diode-pumped erbium-doped microchip laser subject to optical feedback," *Optics letters*, vol. 42, no. 3, pp. 375–378, 2017.

[13] L. Larger, M. C. Soriano, D. Brunner, L. Appeltant, J. M. Gutiérrez, L. Pesquera, C. R. Mirasso, and I. Fischer, "Photonic information processing beyond turing: an optoelectronic implementation of reservoir computing," *Optics express*, vol. 20, no. 3, pp. 3241–3249, 2012.

[14] Y. Paquot, F. Duport, A. Smerieri, J. Dambre, B. Schrauwen, M. Haelterman, and S. Massar, "Optoelectronic reservoir computing," *Scientific reports*, vol. 2, p. 287, 2012.

[15] S. Ortín, M. C. Soriano, L. Pesquera, D. Brunner, D. San-Martín, I. Fischer, C. Mirasso, and J. Gutiérrez, "A unified framework for reservoir computing and extreme learning machines based on a single time-delayed neuron," *Sci. Rep.*, vol. 5, p. 14945, 2015.

[16] F. Duport, A. Smerieri, A. Akrout, M. Haelterman, and S. Massar, "Fully analogue photonic reservoir computer," *Scientific reports*, vol. 6, p. 22381, 2016.

[17] L. Larger, A. Baylón-Fuentes, R. Martinenghi, V. S. Udaltsov, Y. K. Chembo, and M. Jacquot, "High-speed photonic reservoir computing using a time-delay-based architecture: Million words per second classification," *Physical Review X*, vol. 7, no. 1, p. 011015, 2017.

[18] A. Dejonckheere, F. Duport, A. Smerieri, L. Fang, J.-L. Oudar, M. Haelterman, and S. Massar, "All-optical reservoir computer based on saturation of absorption," *Optics express*, vol. 22, no. 9, pp. 10 868–10 881, 2014.

[19] F. Laporte, A. Katumba, J. Dambre, and P. Bienstman, "Numerical demonstration of neuromorphic computing with photonic crystal cavities," *Optics express*, vol. 26, no. 7, pp. 7955–7964, 2018.

[20] K. Vandoorne, P. Mechet, T. Van Vaerenbergh, M. Fiers, G. Morthier, D. Verstraeten, B. Schrauwen, J. Dambre, and P. Bienstman, "Experimental demonstration of reservoir computing on a silicon photonics chip," *Nature communications*, vol. 5, p. 3541, 2014.

[21] F. Denis-Le Coarer, M. Sciamanna, A. Katumba, M. Freiberger, J. Dambre, P. Bienstman, and D. Rontani, "All-optical reservoir computing on a photonic chip using silicon-based ring resonators," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 24, no. 6, pp. 1–8, 2018.

[22] Q. Vinckier, F. Duport, A. Smerieri, K. Vandoorne, P. Bienstman, M. Haelterman, and S. Massar, "High-performance photonic reservoir computer based on a coherently driven passive cavity," *Optica*, vol. 2, no. 5, pp. 438–446, 2015.

[23] G. Van der Sande, D. Brunner, and M. C. Soriano, "Advances in photonic reservoir computing," *Nanophotonics*, vol. 6, no. 3, pp. 561–576, 2017.

[24] M. Lukoševičius and H. Jaeger, "Reservoir computing approaches to recurrent neural network training," *Computer Science Review*, vol. 3, no. 3, pp. 127–149, 2009.

[25] A. Rodan and P. Tino, "Minimum complexity echo state network," *IEEE Trans. Neural Netw.*, vol. 22, no. 1, pp. 131–144, 2011.

[26] L. Appeltant, M. C. Soriano, G. Van der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. R. Mirasso, and I. Fischer, "Information processing using a single dynamical node as complex system," *Nature communications*, vol. 2, p. 468, 2011.

[27] D. Brunner, B. Penkovsky, B. Marquez, M. Jacquot, I. Fischer, and L. Larger, "Tutorial: Photonic neural networks in delay systems," *Journal of Applied Physics*, vol. 124, no. 15, p. 152004, 2018.

[28] M. C. Soriano, J. García-Ojalvo, C. R. Mirasso, and I. Fischer, "Complex photonics: Dynamics and applications of delay-coupled semiconductors lasers," *Rev. Mod. Phys.*, vol. 85, pp. 421–470, Mar 2013.

[29] A. Argyris, D. Syvridis, L. Larger, V. Annovazzi-Lodi, P. Colet, I. Fischer, J. Garcia-Ojalvo, C. R. Mirasso, L. Pesquera, and K. A. Shore, "Chaos-based communications at high bit rates using commercial fibre-optic links," *Nature*, vol. 438, no. 7066, pp. 343–346, Nov 2005.

[30] M. Peil, I. Fischer, W. Elsäßer, S. Baki, N. Damaschke, C. Tropea, S. Stry, and J. Sacher, "Rainbow refractometry with a tailored incoherent semiconductor laser source," *Applied Physics Letters*, vol. 89, no. 9, p. 091106, 2006.

[31] A. Uchida, K. Amano, M. Inoue, K. Hirano, S. Naito, H. Someya, I. Oowada, T. Kurashige, M. Shiki, S. Yoshimori *et al.*, "Fast physi-

cal random bit generation with chaotic semiconductor lasers," *Nature Photonics*, vol. 2, no. 12, p. 728, 2008.
[32] R. Lang and K. Kobayashi, "External optical feedback effects on semiconductor injection laser properties," *IEEE J. Quantum Electron.*, vol. 16, no. 3, pp. 347–355, 1980.
[33] K. Ikeda, "Multiple-valued stationary state and its instability of the transmitted light by a ring cavity system," *Optics communications*, vol. 30, no. 2, pp. 257–261, 1979.
[34] T. Erneux, *Applied delay differential equations*. Springer Science & Business Media, 2009, vol. 3.
[35] M. Peil, M. Jacquot, Y. K. Chembo, L. Larger, and T. Erneux, "Routes to chaos and multiple time scale dynamics in broadband bandpass nonlinear delay electro-optic oscillators," *Physical Review E*, vol. 79, no. 2, p. 026208, 2009.
[36] K. Ikeda and K. Matsumoto, "High-dimensional chaotic behavior in systems with time-delayed feedback," *Physica D*, vol. 29, no. 1, pp. 223 – 235, 1987.
[37] Y. Kuriki, J. Nakayama, K. Takano, and A. Uchida, "Impact of input mask signals on delay-based photonic reservoir computing with semiconductor lasers," *Optics express*, vol. 26, no. 5, pp. 5777–5788, 2018.
[38] K. Hicke, M. A. Escalona-Morán, D. Brunner, M. C. Soriano, I. Fischer, and C. R. Mirasso, "Information processing using transient dynamics of semiconductor lasers subject to delayed feedback," *IEEE J. Sel. Top. Quantum Electron.*, vol. 19, no. 4, pp. 1 501 610–1 501 610, July 2013.
[39] R. M. Nguimdo, G. Verschaffelt, J. Danckaert, and G. Van der Sande, "Reducing the phase sensitivity of laser-based optical reservoir computing systems," *Optics express*, vol. 24, no. 2, pp. 1238–1252, 2016.
[40] J. Vatin, D. Rontani, and M. Sciamanna, "Enhanced performance of a reservoir computer using polarization dynamics in VCSELs," *Optics letters*, vol. 43, no. 18, pp. 4497–4500, 2018.
[41] M. Schmidt, N. Le Roux, and F. Bach, "Minimizing finite sums with the stochastic average gradient," *Mathematical Programming*, vol. 162, no. 1-2, pp. 83–112, 2017.

Apostolos Argyris received the B.Sc. degree in physics from the Aristotle University of Thessaloniki, Greece, in 1999, the M.Sc. degree in Microelectronics & Optoelectronics from the University of Crete, Greece in 2001 and the Ph.D. degree from the Department of Informatics and Telecommunications, National & Kapodistrian University of Athens (NKUA), Greece, in 2006, where he continued working as a research scientist until 2016. He held an Adjunct Lecturer position in the Departments of Computer Engineering, Telecommunications and Networks and Informatics of the University of Thessaly, Greece, from 2006 to 2009 and from 2014 to 2015. In 2016 he joined the Institute for Cross-Disciplinary Physics and Complex Systems - IFISC (CSIC-UIB) in Spain as a Marie-Skłodowska Curie Fellow. He has published more than 80 papers in peer-reviewed journals and conference proceedings, while he is the main author of 3 book chapters. His research interests include nonlinear dynamics in semiconductor lasers, photonic devices and coupled systems, optical communication systems, chaotic encryption, optical sensing and machine learning techniques. Dr. Argyris was distinguished as one of the world's top young innovators from the Technology Review magazine of Massachusetts Institute of Technology, Boston, U.S.A. by receiving the TR35 Award in 2006. The same year he was awarded the Ericsson Telecommunications Award of Excellence for his PhD thesis.

Javier Cantero was born in Palma de Mallorca, Spain, in 1996. He received his bachelor degree in Physics from the University of the Balearic Islands, Spain, in 2018. He currently holds a full-time research position at the company Wireless DNA, where he relates meteorology and telecom systems in a project financed by the Centre for the Development of Industrial Technology (CDTi). His main research interests include non-conventional meteorological observation methods and machine learning, in particular, deep learning based systems.

Marcos Galletero Romero majored in physics at Universidad Complutense de Madrid. Upon its completion, he spent a year at the Institute for Cross-Disciplinary Physics and Complex Systems, where he assisted Miguel Cornelles Soriano and Claudio R. Mirasso in machine learning topics and reservoir computing simulations. He currently works as a data scientist and programmer for various companies on topics ranging from time series analysis to natural language processing and cloud computing.

Ernesto Pereda was born in Madrid on 1973. He got his degree on Applied Physics on 1997 and his Ph. D. in Neuroscience in 2001. He obtained his tenure in Electrical Engineering on 2009 in the Universidad de La Laguna, where he is now a Full Professor.

Claudio R. Mirasso received the Ph.D. in physics from the Universidad Nacional de La Plata, Argentina, in 1989. He has held post-doctoral positions in Spain and the Netherlands. He is Full Professor at the Physics Department, Universitat de les Illes Balears, Palma de Mallorca, Spain and researcher of the Institute for Cross-Disciplinary Physics and Complex Systems, joint center of the Spanish National Research Council and the University of the Balearic Islands. He has authored or co-authored over 150 journal papers. He was coordinator of the European Projects OCCULT and PHOCUS. His current research interests include dynamics of semiconductor lasers, synchronization and control of dynamical systems, dynamics and applications of delayed coupled systems, information processing, neuronal dynamics and applications of nonlinear dynamics in general.

Ingo Fischer received the Diploma and Ph.D. degrees in physics from Philipps University, Marburg, Germany, in 1992 and 1995, respectively. Since 2009, he has been a Professor at the Institute for Cross-Disciplinary Physics and Complex Systems, joint center of the Spanish National Research Council and the University of the Balearic Islands, Palma de Mallorca, Spain. His current research interests include nonlinear photonics and neuro-inspired information processing, and in particular, the emission properties and dynamics of modern photonic sources, coupled laser systems, synchronization of lasers and neurons, information processing in and by complex systems and utilization of chaos.

Miguel Cornelles Soriano was born in Benicarlo, Spain, in 1979. He is member of the IEEE Photonics Society. He received the Ph.D. degree in applied sciences from the Vrije Universiteit Brussel, Brussels, Belgium, in 2006. He currently holds a tenure-track position at the University of the Balearic Islands, Spain. His main research interests cover topics of nonlinear dynamics and information processing based on reservoir computing. As an author or co-author, he has published 55 research papers in international refereed journals.